

# The Popper-Miller theorem is the Bayesian transitivity paradox.

Ryan Giordano

July 21, 2023

Popper and Miller[1,2] proposed a tidy little paradox about inductive reasoning. Many 20th century Bayesians (e.g. [3]) claim that Bayesian reasoning is valid inductive reasoning. Popper, ever the enemy of induction, produced (with Miller) the Popper-Miller (PM) theorem, which "proves" that Bayesian "induction" is nothing but watered-down deduction.

The PM theorem was widely discussed at the time, and I feel there are enough counterarguments. Here, I would like to point out something different: to argue that the PM theorem is just the Bayesian transitivity paradox (BTP) in fancy dress. That there is a connection between the PM theorem and the BTP was pointed out in a brief comment by Redhead in [4], but I think the connection is deeper and simpler than has been noticed before.

I'll first say what the PM theorem and BTP are, and then show that the two are the opposite sides of the same coin. The setup, throughout, is as follows. Suppose that we are interested in how much some observed evidence,  $e$ , supports a hypothesis  $h$ , when  $h \Rightarrow e$ , but  $e \not\Rightarrow h$ . For example,  $h$  might be "my coin has heads on both sides" and  $e$  might be "I observed a heads after a single flip."

I'll use logic notation ( $\vee$  for disjunction,  $\wedge$  for conjunction,  $\neg$  for negation), but one could equally have represented  $e$  and  $h$  as sets than as logical propositions (with  $\cup, \cap, (\cdot)^c$  instead of the respective logic symbols). Note that logical implication of propositions ( $A \Rightarrow B$ ) is the same as set containment ( $A \subseteq B$ ).

Both the PM theorem and the BTP are stated in terms of Bayesian logic. So I'll begin by assuming that I have a measure  $p(\cdot)$  on propositions, where  $p(\cdot|\cdot)$  denotes conditional probability, though my final conclusion will be in much greater generality. In particular,  $p(h|e)$  is the posterior credibility of  $h$  given that we observed  $e$ . Popper and Miller analyzes the "support" of

$e$  for  $h$ , which is defined as the difference between the posterior and prior probabilities of  $h$ , i.e.,  $s(h|e) := p(h|e) - p(h)$ . When support is positive, we say  $e$  supports  $h$ , and when it is negative, we say  $e$  counter-supports  $h$ . We'll assume that  $s(h|e) > 0$  here. I assume throughout that  $p(e)$ ,  $p(h)$ , and  $p(h|e)$  are all strictly between 0 and 1, which is not essential, but simplifies things a bit.

## 1 The Popper-Miller (PM) Theorem

The PM theorem is based on a decomposition of  $h$  into "deductive" and "inductive" parts, denoted  $h_D$  and  $h_I$  respectively, with  $h = h_D \wedge h_I$ . The deductive part has the property that  $e \Rightarrow h$ , and the inductive part is supposed to capture "all of  $h$  that goes beyond  $e$ ." Their particular decomposition doesn't matter for my purposes (it happens to be  $h_D = h \vee e$  and  $h_I = h \vee \neg e$ ), but it has these properties:

$$(A) \quad s(h|e) = s(h_D|e) + s(h_I|e) \quad (1.1)$$

$$(B) \quad s(h_I|e) < 0. \quad (1.2)$$

Property (A) supports the notion that  $h_D$  and  $h_I$  are a "decomposition" of the support of  $e$  for  $h$ .

Property (B) is the PM theorem. It says that the inductive component is always counter-supported by the evidence. One might interpret (B) as follows: that Bayesian reasoning appears to do induction is only an illusion. The support is merely deductive support diluted by inductive counter-support. Or so, at least, Popper and Miller claim.

The most common objection was whether reasonable alternative decompositions exist. Of course they do, and most of the discussion in the literature was about precisely which candidate decompositions are valid and not. Some decompositions violate (A) and some violate (B). Popper and Miller argue in several ways that their decomposition is uniquely appropriate (Popper and Miller, 1987), though I think that Elby and Redhead argue convincingly that other decompositions are reasonable [4,5]. For my purpose, all that matters is that a family of alternative decompositions exist, some of which may violate either (A) or (B).<sup>1</sup>

---

<sup>1</sup>A family of decompositions satisfying  $h = x_D \wedge x_I$ ,  $e \Rightarrow x_D$ , and condition (A) can be found by taking  $e, a, b$  to be any partition of the tautology (so that  $p(e \vee a \vee b) = p(e) + p(a) + p(b) = 1$ ), and taking  $x_D = e \vee a$  and  $x_I = e \vee b$ . Levi pointed out one such

## 2 The Bayesian Transitivity Paradox

A consequence of (A) and (B) which is not much remarked on in the PM theorem debate is the following:

$$(C) \quad s(h_D|e) - s(h|e) > 0.$$

That is, the deductive component receives greater support than the hypothesis. This makes intuitive sense as a desiderata for a decomposition:  $h \Rightarrow e \Rightarrow h_D$ , and, intuitively, any notion of "support" should give no more support to a hypothesis than to its logical consequences.

One might wonder whether it is always the case that  $s(r|e) > s(q|e)$  when  $q \Rightarrow r$ . It turns out that this is not necessarily the case, a phenomenon that is known as the "Bayesian transitivity paradox" (BTP). In fact, the  $h_I$  component of the PM theorem is an example:  $s(h|e) > 0 > s(h_I|e)$ , although  $h \Rightarrow h_I$ . So the PM theorem unavoidably involves the BTP, a point noted by Redhead [5].

It's worth noting that the posterior itself does not suffer from anything like the BTP. If  $q \Rightarrow r$ , then  $p(q|e) \geq p(r|e)$ , since logical implication is the same as set containment. The BTP occurs for  $s(\cdot|e)$  because of the role played by the prior.

Of course, (C) follows from (A) and (B), and (B) follows from (A) and (C). It follows that, given a decomposition of the form (A), *the inductive support is negative if and only if the deductive part has greater support than the original hypothesis*.

## 3 The PM theorem is a special case of the BTP

Let us step back from the specific notion of support and decomposition used in the PM theorem, and ask what we might want *in general* from a decomposition of a generic notion of support, which we denote  $\sigma(\cdot|e)$  into a deductive and inductive part, which we call  $x_D$  and  $x_I$ . We no longer require  $h = x_D \wedge x_I$ , but we do require that  $e \Rightarrow x_D$  in some sense. To investigate a generalized form of the PM theorem, one might ask whether we can have:

---

decomposition in [6], though argued that Bayesian inference is still not deduction since  $s(h_I|e)$  varies over possible decompositions despite the fact that  $h_I \wedge e = h$  for all such decompositions, and propositions that are logically equivalent given  $e$  should receive equal support from  $e$ . Other authors, e.g. Jeffreys in [7], argue for decompositions that violate (A). It is easy to show that if  $e \vee a \vee b$  is not the tautology, then (A) is violated.

$$\begin{array}{ll}
(A') & \sigma(h|e) = \sigma(x_D|e) + \sigma(x_I|e) \\
(B') & \sigma(h_I|e) \geq 0 \\
(C') & \sigma(h_D|e) \geq \sigma(h|e).
\end{array}$$

We want (A') because that's what a "decomposition" would mean, we want (C') because we don't want anything like the BTP, and we want to know whether (B') is possible because that's what it would mean to do induction. But obviously, by basic algebra, (A'), (B') and (C') cannot be simultaneously true, for *any* possibly notion of support, probabilistic or otherwise. The PM theorem is simply a particular case of this simple and general observation.

In light of this, the PM theorem begins to look a little trivial. When Popper and Miller insist, e.g. in response to [7], that authors who contest their decomposition produce alternative decompositions, they are in fact begging the question.

That the BTP occurs is certainly a meaningful critique of probabilistic support  $s(\cdot|\cdot)$ . It seems to me, however, that the PM theorem simply rearranges the BTP in a way that sacrifices clarity rather than illuminates what is really at issue.

## References

Popper, K. and D. Miller (1987). "Why probabilistic support is not inductive". In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 321.1562, pp. 569–591.